

DOCUMENT RESUME

ED 295 984

TM 011 817

AUTHOR Glidden, H. Guy
TITLE The Use of Test Score Growth Measurement as a Positive Basis for School Improvement.
PUB DATE Apr 88
NOTE 13p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 5-9, 1988).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Educational Assessment; Educational Development; *Educational Improvement; Elementary Secondary Education; Evaluation Methods; *Norm Referenced Tests; Predictive Measurement; *School Effectiveness; *Scores; Standardized Tests; *Testing Programs
IDENTIFIERS *Test Score Growth Measurement; *Wichita Public Schools KS

ABSTRACT

During the 1985-86 school year, an evaluation plan utilizing standardized norm-referenced test information was initiated within the Wichita Public School System in Kansas. The plan had four identifiable factors: (1) a growth measurement concept; (2) a simplified computation process; (3) a means of integration of growth measurement within user services; and (4) a means of balancing growth information within the evaluation process. The plan provided evaluation information based on the comparison of actual and predicted growth of test scores resulting from an annual administration of the norm-referenced tests to district students. Two factors that had a major impact on identification of the plan were wide differences of student populations in different schools in the district and an existing annual standardized testing program. Comprehension of the plan by teachers and principals was central to implementation. The impact of the plan on the evaluations of school improvement plans and of the principals has been inconclusive due to administrative changes within the district. The plan, however, has had a positive impact on the effective use of test information by the professional staff of the district. (TJH)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED295984

**THE USE OF TEST SCORE GROWTH MEASUREMENT AS A
POSITIVE BASIS FOR SCHOOL IMPROVEMENT**

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☒ This document has been reproduced as
received from the person or organization
originating it.
☐ Minor changes have been made to improve
reproduction quality.

- Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

H. GUY GLIDDEN

by

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

**H. Guy Glidden, Ph.D., Director
Pupil Evaluation and Testing
Wichita Public Schools**

**Paper Presented at the Annual Meeting of
the American Educational Research Association
April, 1988
New Orleans, Louisiana**

TM011817

The Use of Test Score Growth Measurement as a Positive Basis for School Improvement

H. Guy Glidden, Ph.D.

INTRODUCTION

Notable in recent years has been an increase in the use of standardized norm-referenced test scores as measures of educational quality (Hoover, 1983). The rationale for public expectations that standardized tests are appropriate for use in this manner is understandable. Standardized tests are widely used, relatively inexpensive and provide a ranking format for test results. Because the publishers of norm-referenced standardized tests provide percentile ranks for schools as well as for individual pupils, local decision-makers logically expect that these same test scores could be used to evaluate the work being performed by school personnel or to evaluate the success of educational programs within the schools.

Although demands for evaluations about the quality of educational programs are certainly not new, events of recent years have accentuated demands for comprehensive evaluation systems. The report "A Nation at Risk" increased concerns within school districts about the quality of local schools and placed greater expectations for clear and concise evaluations on the shoulders of local administrators. Further, the publicity gained by the wall chart concept of the U. S. Department of Education for ranking the educational programs of states appears to have reinforced expectations that evaluations of complex educational programs can be reduced to simplified numeric indicators of quality. Again, such expectations underscore the use of test scores as the available means to quantify measures of quality.

By process of elimination, the available tool for evaluation in many school districts is the norm-referenced standardized test. Although examinations which measure the mastery of learner objectives within identified courses or educational programs would appear to be ideally suited for program evaluations, comprehensive programs of criterion-referenced tests are expensive and take time to develop (Northwest Evaluation Association, 1986). For many school boards and chief school officers, it is not possible to justify an increase of evaluation costs for a system which can not produce results for at least three years. On the other hand,

norm-referenced tests are already in place in most school districts, are relatively inexpensive and can produce reliable test results very quickly.

The weaknesses of using information provided by norm-referenced tests are clear. First, standardized tests provide general measures of educational skills. Although test publishers assert that the skills being tested with these tests are in substantial agreement with the educational objectives of most schools (Mehrens and Green, 1986), this assumption must be reviewed with care in each separate school district. If the skills being measured are not consistent with the objectives being taught, test results can not be viewed as valid evaluations of schools or of the work of educators in these schools. Since standardized tests can not be altered to meet local needs, the alternatives--altering the curriculum to suit the test, selecting a different test or simply accepting a mismatch of tests and teaching--may not be desirable. Second, the most easily understood feature of norm-referenced tests is the percentile rank. The percentile rank concept appears to the general public to be the obvious basis for evaluations even though test publishers and educators emphasize that standardized tests measure skills which result from a combination of school and non-school influences. Thus, standardized test scores provide a combination measure which may provide a more accurate view of the supportive resources of parents in the community than of the efforts of school personnel (Lewis, 1986).

THE DEVELOPMENT OF GROWTH MEASUREMENT PLANS

A mandate to use a standardized test as an evaluation tool is feasible if two actions or adjustments are made. First, action must be taken to verify the level of agreement between the skills being tested and the concepts being taught. Although the publishers of standardized tests assert that the skills being assessed by most standardized test batteries are consistent with the educational programs in most school districts, a careful review process is necessary to verify and communicate within a community; a) the level of agreement which exists and b) adjustments which may be necessary either in the use of information or in emphasis which must be placed on teaching certain skills.

Second, an adjustment must be made in the manner of using information from standardized tests as an evaluation tool. In addition to percentile rank information, all standardized tests report a standard score, allowing for a view of year to year growth by individual pupils. Thus, by charting the relationships of standard scores and percentile ranks for individual pupils, it is possible to note gains or losses in relation to all other pupils who shared a common initial score or set of initial scores. In this manner the status score becomes less important than the relative gains made by

test takers. All pupils can be provided for goals for improvement and each can be given a realistic goal based on past measures of achievement. Through the development of a method to combine growth information for school units, an evaluation model using norm-referenced test information can be made both feasible and fair.

The use of growth measurement plans for school evaluation is not a new concept (Azumi, 1987, Gastright, 1987, Webster, 1985). The concept is particularly appealing in school districts where the range of student backgrounds and the difference between school units is very wide. Although the school unit can not be held responsible for the preparation experiences of pupils in attendance, it is reasoned that schools can be evaluated on the basis of the climate provided for academic growth (Webster, 1985). Schools, in order to be successful in a growth measurement system must; a) use existing test information to correctly identify the needs of the pupils in that setting and b) identify a series of teaching and support strategies which will be most effective in meeting these needs. Within such a system, the most successful school in a school district could also be the school with the lowest average test scores.

THE WICHITA PLAN FOR GROWTH MEASUREMENT

Many of the factors which lead to the development of a growth measurement plan based on the use of norm-referenced standardized tests in the Wichita Public Schools are not unique to this setting.

1. Wichita is an urban school district with approximately 46,000 pupils who are being served in over 90 elementary and secondary schools. Although the racial composition of schools is balanced with the use of a cross district busing plan, there is a wide variation of community and income factors among Wichita schools.
2. The retirement of a sixteen year veteran chief school officer in 1984 presented both the opportunity and mandate to review current programs. Subsequently, the new chief school officer placed a high priority on the development of an evaluation system for the school improvement plans by building principals.
3. Rapid changes in expectations for accountability after a long period of stability brought into the open major concerns about the overall development of the evaluation system and about the specific use of test scores as a factor of evaluation.
4. Although a program of annual pupil assessments with the use of standardized norm-referenced tests existed at the time of the mandate for an evaluation system (1985), a coordinated program for the development and use of subject matter criterion-referenced tests did not exist and was not initiated until 1986. Review of the

standardized testing program in the district had consistently shown high levels of agreement between skills tested and stated district objectives in the areas of reading, math and language arts skills.

The plan developed and implemented for evaluation in the Wichita Public Schools has four identifiable factors. These are:

1. the growth measurement concept.
Two successive earned test scores are used to provide a prediction of the succeeding test score for each pupil. A comparison of the earned and predicted scores is identified for each pupil and the relative gains of all pupils in the school unit are combined to provide a school growth index. Since each pupil has a unique predicted score, the task of pupil needs assessment and academic program delivery is different for each school. Also, the composite score is used for comparison purposes so that emphasis within the school will be placed on the development of a total climate for academic growth.
2. a simplified computation process.
The unique feature of the Wichita Growth Measurement Plan is the simplified method of computing gains. Three features are emphasized in a direct effort to increase the level of staff and administrative understanding.
 - a) Grade equivalent scores--long used in local inservice programs on testing and in the release of test scores--were directly used in the computation process.
 - b) The computation process was deliberately simplified. In fact, the computation of a growth prediction for a single pupil can be completed with the use of a hand held calculator and information about the process is available to schools and parents.
 - c) The development of a school growth index begins with a separate computation of growth for each pupil. Thus, each school must focus on the development of an educational climate which is appropriate for the unique pupils in that school.
3. integration of growth measurement within user services.
In order to support the acceptance of the growth measurement concept, information generated by the growth measurement evaluation system is presented in a manner which directs the efforts of administrators and teachers towards the most effective and efficient methods of achieving positive results within the evaluation system. This has been accomplished by:
 - a) designing report forms which clearly present growth information, b) making growth information readily available to school personnel, and c) developing staff information/inservice programs to emphasize effective use of growth report information.

4. balancing growth information within the evaluation process. In order to provide a balanced evaluation process for the job responsibilities of school administrators, two important steps have been taken. Test score growth was intended to represent only 20% of the total evaluation of administrators and was to be balanced with other expectations, and test score growth information was provided to line directors with the understanding that evaluations for school administrators would be based on a review of the total school growth process.

A key element of efforts to gain acceptance of the growth measurement concept in Wichita has been attempts to increase understanding of the concept through the promotion of inservice/information support services to schools, emphasizing the advantages of using all aspects of test information which are made available to schools. It became apparent that Wichita teachers have understandings of; a) the relationship of earned grade equivalent scores with the grade level of pupils tested, and b) the rationale that different amounts of growth (expressed in grade equivalent scores) are realistic for different pupils. With this background of information, it was possible to illustrate the growth measurement process. Two major needs were served by the inservice/information process. First, the focus for test score improvement was placed on support for individual learners. Through the illustration of computing individual gains, it was clearly demonstrated that each individual would have an impact on school totals and that the expectations for individual pupils were realistic in view of demonstrated test score records. Second, the additional view of pupil test scores as provided through the growth measurement process has prompted requests for more test reports which are a) specific in the identification progress in specific skill areas, b) provided at more frequent intervals and c) formatted for more convenient distribution to test information users.

RESULTS OF THE WICHITA GROWTH MEASUREMENT PLAN

1. The growth measurement plan has had no impact on the evaluations of school administrators. This is due to the changes of school leadership for the school district. Soon after the plan was initiated during the 1985-86 school year, the superintendent who had mandated the evaluation plan was dismissed by the board of education and replaced by an interim superintendent until July, 1987. Because the current chief school officer has not completed his initial year in this position, the impact of the plan on the evaluations of administrators is not known at this time.

2. The growth measurement plan has stimulated the demand for inservice/information programs on the topic of using test results to improve instruction. Since the initiation of the growth measurement plan at the beginning of the 1985-86 school year, over 150 information programs for administrators and teachers have been provided on the topic of understanding and using test information. This represents more than a 100% increase in demand for information on this topic. Two reasons have been identified for this increase.
 - a. The growth measurement plan has provided a focus for instruction in basic skill areas and a corresponding focus on the improvement of test scores as a means of verifying the impact of instruction. The growth measurement process "upset the apple cart" as a different view was provided.
 - b. A number of teachers and principals reported to their colleagues that use of test score information was a concept which worked and, in turn, sparked increased interest in this topic within the district.
3. The demand for additional test information products by teachers and principals within the district has increased. Because the Wichita Public Schools scores tests and prints reports within the district, staff demand can and does dictate the design of test reports. As a result of the focus provided by the growth measurement plan, two new test information tools have been requested, designed and are now in use. They are:
 - a) a content performance reporting system which provides specific test skill information for individual students as well as class, school and district summary information.
 - b) an on-line computer test inquiry system which provides a total test history for any pupil in the district upon demand.

Thus, school staff members are becoming more specific and exacting in their demands for information as awareness of test information products and the uses of these products increase.

SUMMARY

During the 1985-86 school year, an evaluation plan which utilized standardized norm-referenced test information was initiated within the Wichita, Kansas Public Schools. The plan provided evaluation information based on the comparison of actual and predicted growth of test scores which result from an annual administration of standardized norm-referenced tests to pupils within the district. The development of this specific plan, identified as the Wichita Plan for Growth Measurement, resulted from a review of local needs and resources. Two factors which had a major impact on the identification of the plan were a) wide differences of student populations in different schools in the district

and b) an existing standardized testing program which provided annual assessments of pupils in the district. Particular emphasis was placed on the development of an evaluation system which could be understood by the teachers and principals in the district. The impact of the Wichita Plan for Growth Measurement on the evaluations of school improvement plans and of the principals who developed these plans has been inconclusive because of administrative changes within the district. The plan, however, has had a positive impact upon the effective use of test information by the professional staff of the school district.

Bibliography

Azumi, Jann E., "Effective Schools Characteristics, School Improvement and School Outcomes: What are the Relationships?", paper presented at the meeting of the American Educational Research Association, Washington, D.C., 1987.

Gastright, Joseph F., "Identifying Merit Schools Using a Community-Based Definition", paper presented at the meeting of the American Educational Research Association, Washington, D.C., 1987.

Hathaway, Walter E. and Worth, David, "A Review of Research: Effective Schools", unpublished paper, Portland, Oregon, 1983.

Hieronimus, A. N., Hoover, H. D. and Lindquist, E. F., ITBS Forms G/H Teacher's Guide, Levels 9-14, The University of Iowa, Iowa City, Iowa, 1986.

Hoover, H. D., "Issues in Testing", address presented at Springfield, Missouri and Wichita, Kansas, 1983.

Lewis, Janice C., "The Relationship of School Background Factors to School Building and School District Achievement", paper presented at the meeting of the American Educational Research Association, San Francisco, California, 1986.

Mehrens, William A. and Green, Davis Ross, "Standardized Tests and School Curricula", unpublished paper distributed by CTB/McGraw-Hill, 1986.

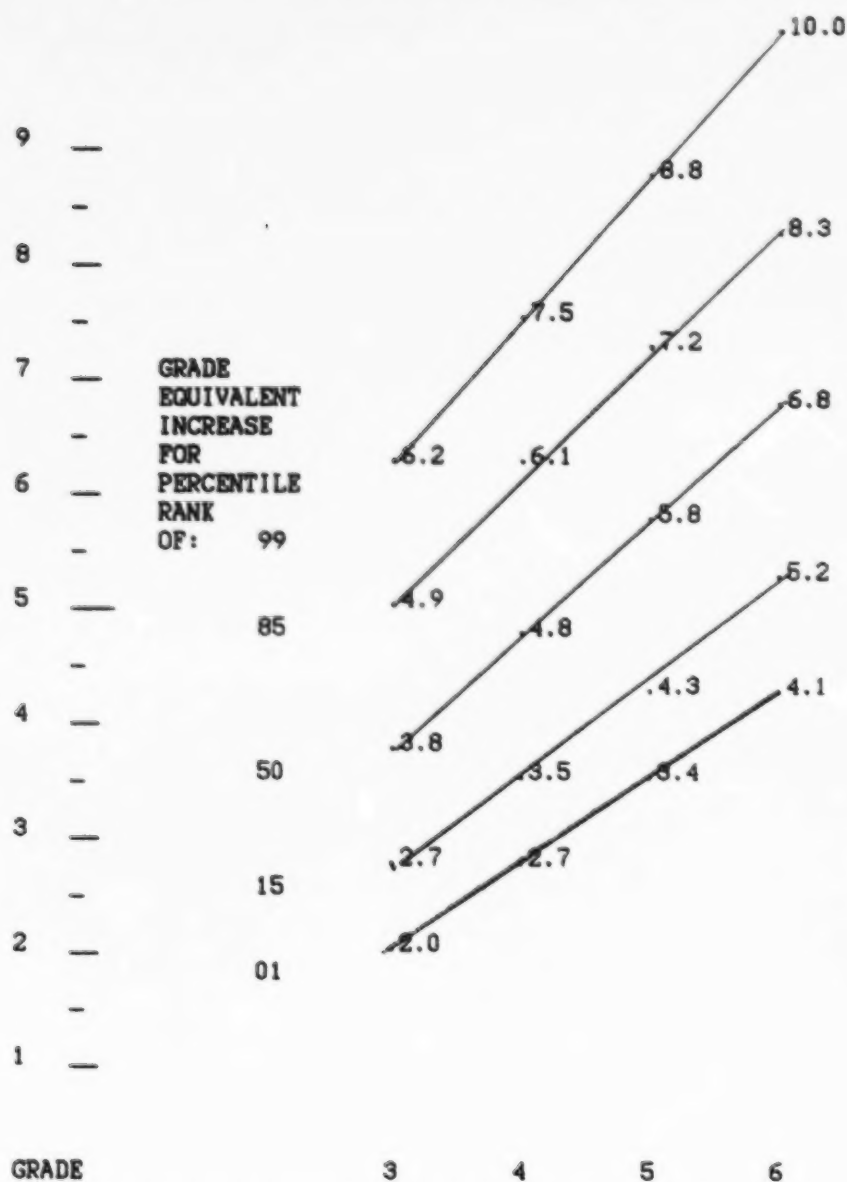
National Commission on Excellence in Education, "A Nation at Risk: The Imperative for Educational Reform", 1982.

Northwest Evaluation Association Curriculum and Assessment Project, "NWEA Trainer's Manual", Salem, Oregon, 1986.

Webster, William J. and Olson, George H., "An Empirical Approach to Identifying Effective Schools", paper presented at the meeting of the American Educational Research Association, New Orleans, Louisiana, 1985.

Illustration #1

ANNUAL GROWTH DIFFERENCES FOR COMPOSITE SCORES
IOWA TESTS OF BASIC SKILLS, 1986 EDITION*



*ITBS Teacher's Guide Levels 9-14, 1986 Edition

Illustration #2

COMPUTATION PROCESS AND EXAMPLES FOR INDIVIDUAL GROWTH GAINS AND SCHOOL GROWTH INDEX WICHITA GROWTH MEASUREMENT PLAN

Computation Process

step #1	earned ratio #1 + earned ratio #2 = combined ratio	where a) earned ratio compares earned composite G.E. and time of test
step #2	combined growth ratio / 2 = ratio	
step #3	growth prediction ratio X test time = predicted score	b) time of test is expressed in G.E. months
step #4	actual score - predicted score = individual growth gain	
step #5	sum of growth gains/ number of test takers = growth index	

Examples

Subj.	Gr 3 comp	Gr 4 comp	Gr 5 comp earned	Gr 5 comp predicted	ind gain
A	4.5	5.7	<u>7.1</u>	6.9	+0.2
B	3.0	5.0	<u>5.6</u>	5.4	+0.2
C	3.8	4.8	<u>5.8</u>	5.8	0.0
D	3.2	3.8	<u>4.4</u>	4.7	-0.0
time of test	3.8	4.8		5.8	
growth index = +0.1/4 (+0.025)					+0.1

individual computation example

- a) $4.5/3.8 + 5.7/4.8 = 2.37$
- b) $2.37/2 = 1.185$
- c) $1.185 \times 5.8 = 6.9$ (predicted growth)
- d) $7.1 - 6.9 = +0.2$ (individual growth gain)

PUPIL EVALUATION AND TESTING
WICHITA PUBLIC SCHOOLS

TEST INFORMATION PRODUCTS FOR
IOWA TESTS OF BASIC SKILLS--KINDERGARTEN-GRADE 8
TESTS OF ACHIEVEMENT AND PROFICIENCY--GRADES 9-11
TEST ADMINISTRATION PERIOD--APRIL

<u>DATE</u>	<u>REPORT</u>	<u>FOR</u>	<u>PURPOSE</u>
early May	class lists school lists skill deficit reports	teachers principals, counselors teachers, principals, subj. coords.	needs ident. progress review planning, ident. for spec. programs
late May	pupil profiles growth reports	parents and pupils teachers, principals, cent. admin.	needs assess. growth ident. review of goals review of instruction
June	content performance summary	principals, subj. coords.	review of instruct. programs, planning for next year
June/ July	media summary	members of media	community information
Sept.	class list reprints* content performance reprints* * fall test reports provide a reorganization of information based on changed grade and location assignments	teachers, principals teachers, principals subj. coords.	information on new pupils development of teaching prescriptions
Oct.	*Profiles of Performance*	board of education, media	detailed review of test info. for schools
available at all times	a) in-depth analysis of test information for specific schools b) test inservice/information programs for teachers, administrators and parents c) individual test histories via on-line computer transactions		